

ABSTRACT

A distributed query engine pipeline architecture comprises cascaded analysis engines that accept an input query and each identifies a portion of the input query that it can pass on to an execution engine. Each stage rewrites the input query to remove the portion identified and replaces it with a placeholder. The rewritten query is forwarded to the next analysis engine in the cascade. Each engine compiles the portion it identified so that an execution engine may process that portion. Execution preferably proceeds from the portion of the query compiled by the last analysis engine. The execution engine corresponding to the last analysis engine executes the query and makes a call to the next higher execution engine in the cascade for data from the preceding portion. The process continues until the results from the input query are fully assembled.